METHODS AND TOOLS FOR THE DEVELOPMENT OF HYDROLOGICALLY CONDITIONED ELEVATION DATA AND DERIVATIVES FOR NATIONAL APPLICATIONS.

Jay R. Kost (jkost@usgs.gov)
Kristine L. Verdin (kverdin@usgs.gov)
Bruce B. Worstell (worstell@usgs.gov)
Raytheon ITSS*
U.S. Geological Survey
EROS Data Center
Sioux Falls, SD 57198

Glenn G. Kelly (kelly@usgs.gov)
U.S. Geological Survey
EROS Data Center
Sioux Falls, SD 57198

Abstract: The National Elevation Dataset (NED) contains the best publicly available elevation data merged into a seamless dataset for the entire United States. In some cases these data contain unwanted artifacts, limiting the quality of standard hydrologic derivatives. The Elevation Derivatives for National Applications (EDNA) project is an interagency effort with the goal of developing a more hydrologically correct version of the NED. This improved NED will be used in the systematic derivation of standard hydrologic derivatives. Methods and tools have recently been developed to facilitate the semiautomatic creation of a hydrologically conditioned NED and hydrologically improved derivatives.

---

*Work performed under U.S. Geological Survey contract 1434-CR-97-CN-40274.

INTRODUCTION

Elevation data have become an increasingly important type of spatial data for applications ranging from visualization to hydrologic modeling. The National Elevation Dataset (NED) has helped many users reduce preparation time by providing elevation data in a pre-processed form. The seamless, edge matched, and filtered nature of the NED makes it a convenient source of elevation data and provides the best available, standard U.S. Geological Survey data for the user community (Gesch, et. al., 2002).

Although the NED is continually updated to include higher quality and higher resolution data, the NED dataset may not meet the hydrologic accuracy requirements of some applications. The goals of the Elevation Derivatives for National Applications (EDNA) project are to create a more hydrologically correct version of NED and to use the conditioned dataset to create a standard hydrologic derivative database. The EDNA database will provide elevation data more suitable for hydrologic applications and will eliminate the pre-processing previously required to generate standard derivatives.

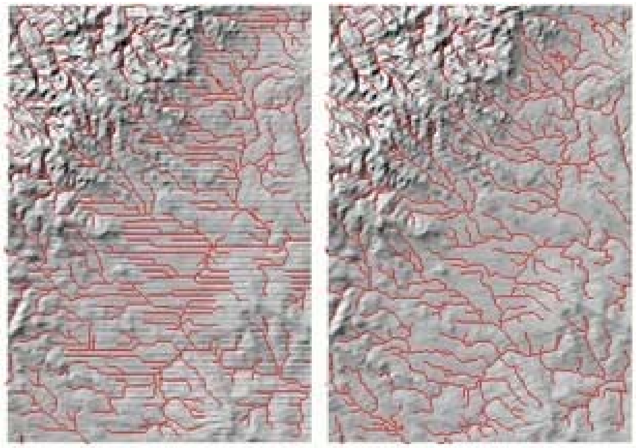
Methods and tools have been developed that make it possible to create hydrologically conditioned elevation data from the NED. This paper provides examples of problems encountered with the NED during the EDNA conditioning process and illustrates the theory behind and use of these methods and tools.

NATIONAL ELEVATION DATASET (NED)

The National Elevation Dataset (NED) is a seamless, digital elevation model (DEM) with coverage of the entire United States, Puerto Rico, U.S. Virgin Islands, and U.S. territories. The NED was developed from individual 7.5-minute DEM's produced by the USGS with the goals of improving data access and minimizing the need for pre-application processing. Figure 1 shows a shaded-relief rendering of the NED for the conterminous U.S.

Figure 1. Shaded relief rendering of the NED.

The NED is updated bimonthly to provide the best publicly available data to the user community. It is constructed with a consistent datum, elevation unit, and projection. NED data are also edge matched and filtered (Oimoen, 2000) to ensure seamless continuity and minimize artifacts. Figure 2 illustrates the flow patterns across a DEM surface before and after the filtering process.



Before After

Figure 2. Flow across a DEM surface can be greatly improved after edge matching and filtering.

Currently, the conterminous U.S., Hawaii, Puerto Rico, and Virgin Islands are distributed with a one-arc-second cell size (approximately 30 meters) while Alaska is distributed with a two-arc-second cell size. Future plans include the incorporation of new data sources that will further increase data quality and resolution. More detailed information about the NED (http://gisdata.usgs.net/ned/) can be found online.

ELEVATION DERIVATIVES FOR NATIONAL APPLICATIONS (EDNA)

The Elevation Derivatives for National Applications (EDNA, http://gisdata.usgs.net/edna/) project is a multi-agency effort that takes advantage of the seamless and filtered characteristics of the NED. The goals of the project are to create a hydrologically conditioned NED dataset, systematically create hydrologic derivatives, and vertically integrate these data with other spatial datasets such as the National Hydrography Dataset (NHD, http://nhd.usgs.gov/).

EDNA data development occurs in three stages, the first of which uses semi-automated techniques to create preliminary derivative data from the unconditioned NED elevation data. The second stage uses the Stage 1 data, within a set of ArcView tools, to create preliminary watersheds and subwatersheds and to identify and flag discrepancies between the Stage 1 derivative data and existing data sets that portray true hydrologic features. Data and information created in Stage 2 are used in Stage 3 to develop a hydrologically conditioned NED dataset. These data are then used to create a more hydrologically correct derivative database. The creation of accurate hydrologic derivatives will facilitate vertical integration of the final EDNA data with the NHD and other datasets.

Stage 1 processing has been completed by the National Weather Services' National Severe Storms Laboratory and the USGS. Processing utilized well tested GIS techniques to create derivative data, including initial synthetic streams and reach catchments (drainage areas corresponding to each stream in the synthetic streamline coverage). Stage 1 processing was completed during the spring of 2002.

Data processing for Stage 2 utilizes the local expertise of many cooperators across the country to help identify and delineate watershed and subwatershed boundaries (Kost and Kelly, 2001). Local expertise is also an important component in the discrepancy identification and flagging process. Cooperators include Federal, State, and local agencies, private organizations, and academia. The watershed and subwatershed delineations created in Stage 2 are used by many of these cooperators for various applications and will be used in a Federal, multi-agency effort to develop the Watershed Boundary Dataset.

Stage 3 processing is directed at utilizing the flagging information from stage 2 and recently developed algorithms to create a hydrologically conditioned DEM dataset. Tool development has been comprised of creating algorithms for quickly and efficiently modifying the DEM data so flow derivatives more accurately represent natural surface flow. Utilities designed thus far utilize various methods for modifying elevation pixel values in specific, limited areas of discrepancy.

ERROR IDENTIFICATION AND FLAGGING

Discrepancy identification is required prior to DEM conditioning. Currently there are two primary methods used for flagging discrepancies. One of these utilizes algorithms that automatically identify possible problems by using distance threshold and intersect criteria. The other uses down stream traces from NHD end points and synthetic catchment/NHD stream fit criteria to help in identifying areas of discrepancy (Worstell, 2001).

Discrepancy Examples: Due to data and resource limitations it is neither possible to find nor practical to address and fix all discrepancies that may exist between the DEM data and reference data sources. Because of this, identifying stream or ridgeline discrepancies of sufficient magnitude or importance to require resolution can be a very subjective task. Algorithms have been developed to help reduce the degree of subjectivity involved in the flagging and subsequent DEM conditioning process. The following sections discuss these algorithms. Figures 3 and 4 illustrate some of the common types of problems encountered during the process.

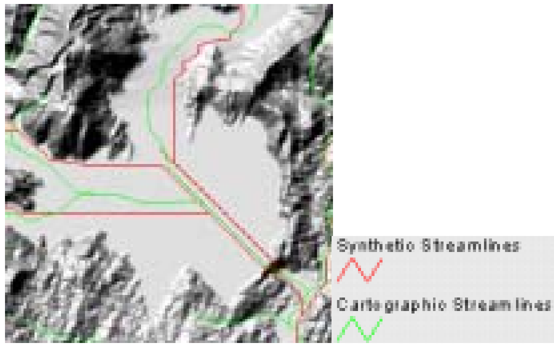


Figure 3. Within flat or low relief areas, elevation data does not have the accuracy or precision necessary to determine proper flow direction (in part due to the need to "fill" sinks). Using Stage 3 tools it is possible to modify the flow to mimic NHD centerlines or other desired flow paths.

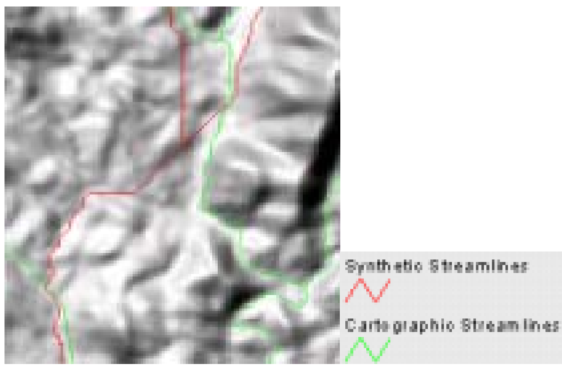


Figure 4. Inaccurate ridge elevations affect both stream flow and watershed boundaries.

Automated Discrepancy Identification: In order to reduce some of the subjectivity and potential to over look problems in the discrepancy flagging process, algorithms were developed and incorporated in both the Stage 2 and Stage 3 tools. Both of the tool sets were developed with ESRI's Avenue programming language and are automatically run by depressing a single user interface button.

The discrepancy identification algorithm uses EDNA, NHD, and existing hydrologic unit boundary data in the identification process. Four different types of problems are identified and layers are created for each so that the Stage 2 analyst can examine each area and create flags as necessary. Three of the problem layers include; arcs indicating NHD and subbasin intersections, synthetic stream and subbasin intersections, and NHD and subwatershed boundary (developed in Stage 2) intersections. Boundary intersections can help identify possible problems with ridgelines. The fourth layer indicates where synthetic and NHD streams are separated by a distance greater than a specified threshold (i.e. 90 meters), indicating possible flow path problems. While the discrepancy layers help to identify problems, it is still up to the analyst to verify the correct flow and document the problem so that it can be corrected in the Stage 3 conditioning process. Local knowledge and ancillary data are used in the verification process.

NHD Downstream Trace: Another means of identifying discrepancies is by using a synthetic stream network and catchments that are developed from NHD stream endpoints. Standard EDNA synthetic stream networks are defined using a constant contributing area threshold to distinguish overland flow from stream flow. The NHD based dataset provides a cartographic representation of stream networks. A downstream trace from the NHD endpoints creates a new synthetic network that more closely mimics the NHD dataset, but the two datasets have varying degrees of spatial agreement.

Catchments based on the new synthetic network are used to match and measure agreement between the EDNA and NHD stream networks. Ideally, there is a one to one correspondence between the two networks. The reality is that confluences may not agree or a synthetic path may significantly deviate from its NHD counterpart. A simple overlay analysis is used to determine NHD reach membership in each catchment and to identify flow path discrepancies. Figure 5 provides an example of where an NHD downstream trace deviates from the cartographic network stream.

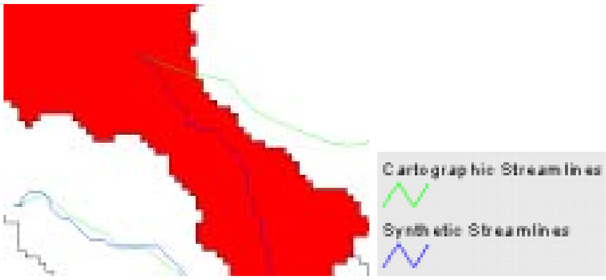


Figure 5. NHD downstream trace helps to indicate where the synthetic flow may be incorrect and subsequently cause the creation of an incorrect catchment boundary.

Polygon And Line Flag Layers: Once the problem identification layers have been generated, the stage 2 analyst can use the information, as well as ancillary datasets (e.g. digital raster graphic maps, digital orthophoto quadrangles, etc.), to quickly move into a specific area and verify if a significant problem exists. If a problem exists (be it with the synthetic stream, boundary, subbasin boundary, or NHD) the analyst would flag it using a flagging tool. This tool allows the analyst to create an attributed polygon around the area. Attributes can then be added, through a pop-up form, which describe the problem and the required solution. A line flag can also be generated which can be used to draw proper flow paths or boundaries when other ancillary data do not exist or are incorrect.

HYDROLOGIC CONDITIONING TOOLS

The hydrologic conditioning tools were developed using Avenue within ArcView. The Spatial Analyst extension is required in order to use these tools. ArcView was chosen for tool development because of the desire to have a developer and user-friendly interface. Future work will be geared toward moving current and future conditioning tools into the ArcGIS, VBA (Visual Basic for Applications) environment and in further automating the editing process.

<u>Tool Organization and Documentation:</u> DEM data, no matter the underlying topography or data quality, inevitably contains problems, which affect the extraction of hydrologic derivatives. The Stage 3 tools were developed to address each of the most common types of problems and were subsequently organized according to the problem they help to address. Organization includes separate view categories for each of the tool types. As each of these views is made active and populated with data, the associated buttons and tools, unique to the specific problem, are activated for use. Figure 6 shows the different view categories.

Figure 6. DEM conditioning views and tools are organized by function.

Documentation has also been developed and is included in the ArcView help system. Topics are available for everything from the general organization of the tools to descriptions of each tool, including examples of where it might be used.

<u>Data Requirements:</u> Initial development of the DEM conditioning tools has been within the context of the EDNA project. For this reason they have been developed to work with specific datasets related to EDNA. While the underlying functionality of the tools can work with any grid source, at present EDNA related datasets are required. Table 1 lists the data used within the tools.

Table 1. Required data for use with editing tools.

Name	File Name	Description
Synthetic Streamlines	streams.shp (line)	DEM derived streamlines
Cartographic Streamlines	nhd_rch.shp (line)	NHD reaches
Original DEM	orig_demf (grid)	Stage 1 DEM (unfilled)
Filled DEM	filled_demf (grid)	Stage 1 DEM (filled)
Shaded Relief	sh_relieff (grid)	DEM shaded relief
Sinks Grid	sinksf (grid)	Sinks filled in fill process
Flow Directions	flow_dirf (grid)	8-direction flow from filled dem
Flow Accumulations	flow_accf (grid)	Flow accumulation from filled dem
Watersheds	10dig_watersheds.shp (polygon)	DEM derived watersheds (Stage 2)
Subwatersheds	12dig_watersheds.shp (polygon)	DEM derived sub-watersheds (Stage 2)
Watershed Seeds	10dig_seeds.shp (point)	Watershed pour points (Stage 2)
Subwatershed Seeds	12dig_seeds.shp (point)	Subwatershed pour points (Stage 2)
Flag Polys	flag_polys.shp (polygon)	DEM error flags (Stage 2)
Flag Lines	flag_lines.shp (line)	DEM error lines (Stage2)
250K HUC	huc_poly (polygon)	Existing HUC data

<u>Data Load And Tool Selection:</u> If all required data are in a single directory it can be loaded automatically into a "Stage 3 Data View" by depressing a single button. The new view is the starting point for all subsequent editing operations. From this view three operations can be performed. These include creation of "Flagging", "Watersheds", or editing views determined by which tool is selected.

Flagging View: Data required for automated discrepancy identification is loaded after this view is created. By making this view active, the user can run the automatic problem identification algorithm described earlier. The created discrepancy layers can then be readily loaded into any desired editing view.

Watersheds View: After a new, hydrologically conditioned DEM has been created it becomes possible to regenerate watersheds and subwatersheds. New delineations can be created by using the stage 2 seeds (points representing stage 2 watershed outlets) edited to match the hydrologically conditioned DEM flow accumulations. Watersheds can also be generated by using delineation data from different sources than EDNA. Individual basins can be created as well by clicking on any flow accumulation (stream) cell.

Editing Views: Hydrologic conditioning is initiated by specifying an extent, within the Stage 3 Data View, that will be the local area of analysis. By selecting a smaller area, processing time is substantially reduced and data management is less cumbersome. Once an extent is specified the user is asked to select the desired tool. After the tool is selected a new view is created within the specific tool category (Cell Edits, Sinks, Obstructs, or Flats) and the required data are loaded. A polygon shapefile is also created in the Stage 3 Data View, which indicates the tool and the number of the editing view created (e.g. C1 for Cell Editing view number 1).

<u>Tool Theory And Usage:</u> DEM data inherently have problems accurately depicting flow in certain situations. The most predominant types of problems deal with obstructions, low relief areas, and sinks. The discussion below discusses the tools and how they can be applied to correct problems associated with these data issues.

Obstruction Clearing: Highway overpasses, dams, and other obstructions can cause flow across a DEM surface to be blocked from following the natural flow path. In these cases, it is desirable to create a path through the obstruction to allow for a proper flow path. The purpose of this tool is to modify the DEM such that the proper flow pattern will be maintained through the obstructed area. Figure 7 shows an example where an overpass has caused flow to follow an erroneous path.

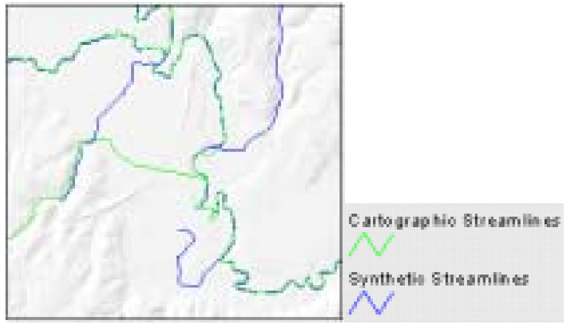


Figure 7. Example area where overpass elevations have caused incorrect flow.

The theory behind this tools operation is based on flow enforcement and iterative elevation modifications. The user first selects the cartographic stream that represents natural flow. The stream is then used to identify underlying grid cells for subsequent processing. These grid cells are first set to "no data", which serves to enforce flow in the grid toward the cartographic stream. The second processing step involves iterating through the original underlying cell values, from upstream to downstream, and modifying the value when necessary to ensure a monotonically decreasing elevation. This step serves to "cut" through any obstruction that may be impeding flow. The final step is to merge the new flow path elevation values back into the original DEM. Figure 8 shows the new DEM after the patch grid creation and merge back into the original. New synthetic streamlines are also shown compared to the cartographic representation.

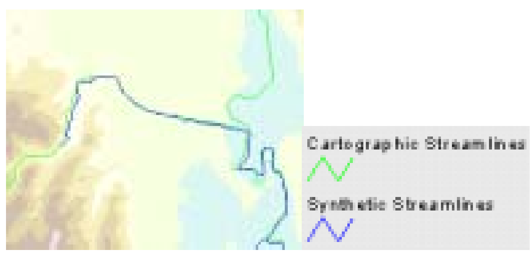


Figure 8. Resulting streamlines from the new DEM.

Flat Modification: Typically, in areas of low relief, neither the accuracy nor the resolution of the existing digital elevation data is adequate to accurately delineate the streamlines that pass through the relatively flat areas. However, in these same areas, the NHD usually contains a reasonably accurate delineation of the streamlines. The objective of using the flats tool is to adjust the flow direction grid to improve the flow so that it more closely matches flow as represented by the NHD. Since this tool deals with the flow direction grid and not the original elevation grid, its use is limited to providing a new flow direction grid only and does not relate directly back to the original elevation data.

The modified flow direction grid creation process begins by utilizing a selected cartographic stream to define the desired flow path. A new grid is then created with the original elevation values maintained underneath the cartographic stream and artificially high elevations elsewhere.

This ensures flow directions toward the cartographic stream. A new flow direction grid is then created, which produces synthetic streams along the cartographic stream path. A flow direction patch can then be created containing the modified flow direction grid cells along the desired flow path. This patch can then be merged back into the original flow direction grid. Figure 9 illustrates the effect that a modified flow direction grid has on synthetic stream generation.

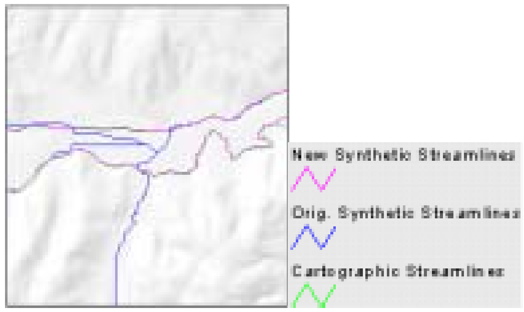


Figure 9. Flow direction modifications can work well for creating new synthetic streams that more closely match cartographic representations.

Sink Clearing: Source data (e.g. map contours) and production methods sometimes cause imperfections to be translated to DEM data. These errors sometimes take the form of spurious "sinks", which can cause problems during generation of flow direction and accumulation grids. In order to create continuous flow, DEM data have to be "filled" which can affect both "real" and "spurious" sinks. The primary problem with the filling process is that real sink features can be filled and low lying topographic detail is sometimes lost. The objective of this tool is to "clear", rather than fill, these sinks so that more topographic detail is preserved in the low areas while continuous flow is maintained.

After defining an extent containing a problem sink or sinks and creating a new edit view, the sink clearing process can be initiated. The clearing algorithm uses an iterative process by which the elevation of the controlling pixel of the sink, in the original grid, are modified to the minimum elevation of its eight neighbors. This will allow that pixel to drain to its neighbor once it has been filled again. The algorithm iterates until the area of the sink has been reduced to a specified number of pixels (e.g. 100). The smaller the desired sink area the more iterations and subsequently the more processing time required. After the cleared grid has been generated a patch grid can be created that contains the modified grid cells. Figure 10 illustrates the effects of having a cleared sink on newly generated streams.

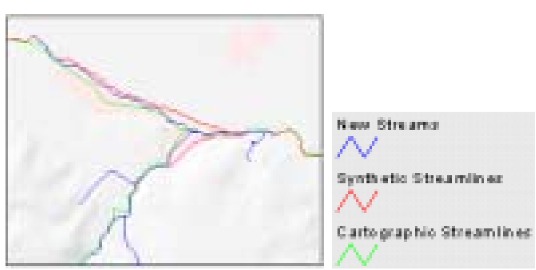


Figure 10. Effects on stream generation after performing sink clearing.

Cell Editing: The cell editing tool is comprised of three different sub categories of tools. All are accessible once a cell editing view is created. The first is a simple, single cell editing tool; the second is a streamline based editing tool; and the third uses a focal statistics algorithm to change underlying cell values beneath selected cartographic streamlines.

If it is determined that only a few cells need to be modified, the single cell editing tool should be used. The tool includes a button for querying elevation values from an active grid. This tool is useful for identifying cells that may have to be modified. Once cells have been identified that need to be edited, the user selects the cell edit tool button. When the user now clicks on a cell to be modified a dialog appears showing the original DEM value and allows the user to enter a new value. When the value is accepted a point is created and added to a generated point shapefile. As additional cells are clicked on for modification, more points are added to the shapefile. The shapefile helps to identify which cells will be modified while still giving the user the ability to finalize the changes prior to actually creating a new grid.

Once all points have been created and accepted as desired cell edits, the point-to-grid button can be depressed. This button converts the points into a new grid with the new elevation values. Once this grid is created it can be merged back into an existing DEM grid. New streams and shaded relief can then be created to verify that the results of the edit(s) are as expected.

The streamline editing tool is similar to the single cell editing tool except that it uses a user specified cartographic stream segment to create points. It also gives the user the option of entering a new elevation value for the entire segment or dropping the elevation value(s) by some specified amount.

The third cell editing tool is similar to the streamline editing tool in that it uses a user selected streamline or streamlines as input for creating point shapefiles that are subsequently converted to a new patch grid. The primary difference with this tool is in the way that the new cell values are generated.

This tool modifies the selected cartographic streamlines underlying cell values by replacing them with those generated within a focal statistics grid. This grid is created by finding the minimum cell value of the original DEM that lies within a block neighborhood, which is specified by the user, and putting this value into a generated streamline grid. Figure 11 shows an example showing the selected NHD stream (red), the original synthetic streamline (cyan), and the new synthetic streamline (blue) generated from the focal stats grid merged with the original DEM.

Figure 11. Resulting synthetic streamline (blue) after using the focal stats cell editing tool.

SUMMARY AND CONCLUSIONS

Elevation data has an important role in topographic science investigations. The NED offers a source of elevation data that provides a consistent datum, elevation unit, and projection and frees the user from having to perform many pre-application processing tasks.

While the NED is a very useful dataset, in some cases the hydrologic accuracy may not meet certain user requirements. In these cases, it becomes desirable to have an elevation dataset that can more accurately depict surface flow. The goals of the EDNA project are to create a hydrologically conditioned elevation dataset and derivatives that can easily be implemented in applications requiring more hydrologically correct elevation data.

An ongoing activity within the EDNA project has been to develop the methods and create the tools needed to edit or condition the elevation data included in the EDNA database. Many tools have been developed that allow the user to create elevation edits, which can be merged back into the original elevation data to create a hydrologically conditioned dataset.

Future work within the EDNA project will include refinement of existing tools as well as the development of new ones. It is also anticipated that efforts will be made toward further automating the editing process. If further automation is successful, it will allow for the quicker development of the EDNA database while maintaining the highest degree of elevation information integrity.

REFERENCES

Gesch, D.B., Oimoen, M.J., Greenlee, S.K., Nelson, C.A., Steuck, M., Tyler, D.J., 2002, The National Elevation Dataset. Photogrammetric Engineering and Remote Sensing, vol. 63, No. 1.

Kost, J.R. and Kelly, G.G., 2001, Watershed Delineation Using The National Elevation Dataset And Semiautomated Techniques, Proceedings of the 2001 ESRI User Conference, July 9-13, 2001, San Diego, CA.

Oimoen, M.J., 2000, An Effective Filter For Removal Of Production Artifacts In U.S. Geological Survey 7.5-Minute Digital Elevation Models, Proceedings of the Fourteenth International Conference on Applied Geologic Remote Sensing, 6-8 November, Las Vegas, NV.

Worstell, B.B, Greenlee, S.K., Verdin, K.L., 2001, Scaling and Integration of Geospatial Topographic Indicators for Hydrologic Applications – Literature Review, U.S. Geological Survey, Open-File Report 2001.